INSTITUTE FOR DEFENSE ANALYSES

Complexity in an Unexpected Place: Quantities in Selected Acquisition Reports

Gregory A. Davis, Project Leader
Margaret L. Giles
David M. Tate

December 2017
Approved for public release;
distribution is unlimited.
IDA Paper P-8490
Log: H 17-000283

INSTITUTE FOR DEFENSE ANALYSES
4850 Mark Center Drive
Alexandria, Virginia 22311-1882

About this Publication
This work was conducted by the Institute for Defense Analyses (IDA) under contract HQ0034-14-D-0001, project number AY-7-4166, "Portfolio Impacts of Program Schedule Changes," for the Director, Performance Assessments and Root Cause Analyses. The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Acknowledgments
Thank you to Kathleen M. Conley, Zachary S. McGregor-Dorsey, and Brian Q. Rieksts for performing technical review of this document.

For More Information:
Gregory A. Davis, Project Leader
gdavis@ida.org, (703) 575-4698

David J. Nicholls, Director, Cost Analysis and Research Division
dnicholl@ida.org, (703) 575-4991

INSTITUTE FOR DEFENSE ANALYSES

IDA Paper P-8490

Complexity in an Unexpected Place: Quantities in Selected Acquisition Reports

Gregory A. Davis, Project Leader
Margaret L. Giles
David M. Tate

Executive Summary

The Department of Defense (DoD) annually produces Selected Acquisition Reports (SARs) for all of their Major Defense Acquisition Programs (MDAPs) and submits these to the Congress. They also tell analysts that this is the primary data source for studying MDAPs. This paper looks at how the SARs report quantities (and the associated unit costs). We examine many examples in which the definition of a unit is non-intuitive, inconsistent, or both. For example, while it might seem straightforward to simply count CH-47 Chinook helicopters produced, there are actually four different variants of CH-47 helicopters within the current Army program. For each of these variants, units produced 10 years ago are significantly different from new units coming off the line today. The CH-47 program is not unusual in showing this kind of unit variation; we have found very few programs in which counting quantities is a simple as the commonly used learning curve models assume it to be.

After discussing the various kinds of unit reporting issues found in recent programs, we also discuss the potential importance of these issues for analysis and oversight activities, and propose ways that unit definitions and reporting might be improved.

Selected Acquisition Reports

The SAR dataset has many appealing characteristics. It reports funds from all different appropriations related to an acquisition program in one place, whereas in the budget submissions these are scattered throughout different exhibits and Services. Each SAR also reports all funding in both base year (BY) and then year (TY) dollars, and quantities from the beginning of the program until its planned conclusion. These reports are used for the calculation of cost growth and for many other types of analysis of the acquisition system.

Quantity Reporting is Not Simple

In most analyses of these data, it is assumed that each unit is essentially identical to every other unit. This is often not the case. In this review, we distinguish three reasons for differences: changes over time, mixed types, and reporting accidents. We found that significant changes over time and mix type issues are not merely common for MDAPs; they are nearly universal. We also found instances of significant accidents in SAR reporting of MDAPs. Few programs are entirely devoid of these issues and many show more than one.

Changes over time refers to cases in which the Service changes the design of the items being purchased from lot to lot, but not the designation. For example, ships in the same class often differ considerably in both configuration and mission systems. Less obviously, successive lots of tactical aircraft or ground vehicles often differ in important ways that affect unit cost.

Mixed types refers to situations in which the program is purchasing multiple distinct end items but does not distinguish among them when counting units. A notable example of this is the Navy's Integrated Defensive Electronic Countermeasures program (IDECM), which purchases both onboard electronics suites and towed decoys and counts each as a unit.

The last reason is *accidents*. While the other two reasons reflect decisions that are made about what a program will buy and how it will report, accidents are cases in which the reported data are incorrect, despite the quality control processes designed to prevent this. Accidents are inherently difficult to find; spotting them requires either knowing the ground truth or recognizing inconsistencies in separate reports that are not designed to be easily reconciled. We found the three instances of accidents by comparing SAR filings and President's Budget justification books.

The prevalence of all of these issues poses serious challenges, both to researchers attempting to understand the causes and mechanisms of cost growth and to oversight bodies attempting to understand cost and capability changes in active programs.

Challenges for Researchers

Researchers have sought to develop predictive models of cost growth, with limited success. McNicol[1] found that reported unit cost growth in MDAPs is closely associated with periods of relatively generous defense budgets. Our findings suggest one possible mechanism for this association—namely, that generous budgets permit programs to add features and correct defects over time, so that units produced in later lots are more capable than those produced in earlier lots. There has been little effort to capture this effect with predictive models, perhaps because few researchers were aware of the need.

Similarly, acquisition researchers have long been interested in trying to predict the effect of changes in production rate on production costs. Schedule instability and production stretch-outs have long been cited as important causes of unit cost growth, but causal models of this effect have been elusive. Our findings suggest that this may in part be due to unaccounted-for variation in the content of units being procured within a given program.

1 David L. McNicol, "Post-Milestone B Funding Climate and Cost Growth in Major Defense Acquisition Programs," IDA Paper P-8091 (Alexandria, VA: Institute for Defense Analyses, March 2017).

Ongoing content change also has implications for how cost analysts should model learning curves. Traditional learning curve theory assumes that the content of all units is identical, but that unit costs decline exponentially as a function of cumulative units produced. Our findings suggest that this is a poor model. For one thing, it fails to account for increasing content (and cost) from lot to lot. For another, it fails to account for losses of learning due to significant design changes. It may be that the "learning and forgetting" model proposed by Benkard[2] works fairly well precisely because the "forgetting" portion of the model can approximately account for both of these effects. Finally, of course, if the units being produced are actually a mix of several different designs, it is difficult to say how much production of one type of unit will drive learning for the other types.

Challenges for Oversight

The acquisition oversight community, from the Congress to the Secretary of Defense down to individual program managers, needs to be able to estimate the likely impacts of changes to a program's acquisition strategy. All of these stakeholders are aware that cost growth is a significant problem for long-term planning and budgeting, but no general-purpose predictive models for program cost growth have been identified. The Congressional Budget Office applies generic cost growth factors for each category of weapon system (such as surface ship, tactical aircraft, or automated information system), but those factors are at best correct on average. Actual cost growth varies widely within each category.

Our findings suggest that some of this cost growth is deliberate, but poorly reported to at least some of the oversight stakeholders. This suggests the possibility of both predictive modeling of that portion of cost growth and improved reporting procedures to inform the Congress and OSD of (possibly contingent) plans for future content change within programs.

Potential Improvements

We propose a few possible adjustments to the reporting system that could make the data more useful both for oversight and analysis. The primary concern is to avoid creating new destructive incentives for acquisition officials. Any change should also avoid placing too much extra burden on staff or revealing too much information to our adversaries around the world.

---

2 C. Lanier Benkard, "Learning and Forgetting: The Dynamics of Aircraft Production," *American Economic Review* 90, No. 4 (September 2000): 1034–1054, http://aida.econ.yale.edu/~lanierb/research/Learning_and_Forgetting_AER.pdf.

Different reasons for unit inconsistency call for different solutions. *Changes over time* calls for data that allow analysts to get a sense of how much units have changed. *Multiple types* calls for a better explanation of what a program is buying. Lastly, reducing *accidents* calls for an examination of the process used to generate SAR data and budget justification submissions.

Contents

1. Introduction

The Director of Performance Assessments and Root Cause Analyses (D,PARCA),[1] asked the Institute for Defense Analyses (IDA) to review the quality and utility of data used for acquisition oversight. This report addresses one particular key facet of acquisition reporting: the definition of units and quantities for major defense acquisition programs (MDAPs). MDAPs are the largest weapon system acquisition programs, and all must report their status annually to the Congress.

Acquisition data are primarily about a few questions: “How much funding?,” “How much are we getting?,” “When are we obligating the funds?,” and, occasionally, “When are we getting what we paid for?” All of these questions are interesting and none are straightforward. Most have been addressed elsewhere and continue to get attention. The question of “What we are getting?,” however, is generally treated in acquisition as though it were simple and not worth in-depth analysis. Our experience tells us that counting quantities is often not straightforward. This report describes the findings of research that has taken us deeper into this question, showing that quantities are almost always complicated.

A. Selected Acquisition Reports

Section 2432 of title 10 U.S.C. requires the Secretary of Defense to submit to Congress a yearly status report for each MDAP.[2] This report, known as the Selected Acquisition Report (SAR), provides comprehensive performance, schedule, and cost data for each MDAP, and is prepared annually in conjunction with the President’s Budget (PB) Submission. In addition to the annual (December) SAR, quarterly SARs are submitted in exceptional cases, such as new program initiation, a cost or schedule breach, or re-baselining of an existing program. Each SAR includes separate cost estimates for research and development, procurement, military construction, and acquisition-related operations

---

[1] PARCA is an office under the aegis of the Office of the Under Secretary of Defense for Acquisition, Technology and Logistics (OUSD(AT&L)).

[2] 10 U.S.C. § 2432: Selected Acquisition Reports, accessed March 28, 2017, http://uscode.house.gov /view.xhtml?req=granuleid:USC-prelim-title10-section2432&num=0&edition=prelim.

and maintenance (O&M). Both past actual costs and future anticipated costs are reported, for the anticipated life of the program.[3]

In addition to costs, the SARs also include the quantity of units purchased each year, and, like funding, they cover the time from the inception of the program until the projected end. Each unit is supposed to be a "fully configured end item."

Program managers (PMs), Program Executive Officers, and/or Service Acquisition Executives use the Defense Acquisition Management Information Retrieval (DAMIR) system to create and update SAR data.[4] Within DAMIR, the "Track to Budget" data section identifies the budget program elements (PEs) and procurement line item numbers (LINs) for each appropriation associated with a program in a particular fiscal year, which allows the user to find the equivalent cost and quantity data in the PB Submission prepared in the same year.[5] Reconciling SAR data with the equivalent PB Submission proves difficult, however, as cost estimates tend to vary between the two sources, and some PEs and LINs are shared among multiple programs in a non-transparent way. In some cases, the SAR and PB submission also measure quantities differently.

Neither the PB nor the SAR is perfect. In general, the Justification Books that the Services produce annually to support the PB, contain more detail, which is good for analysis, but if it extends beyond the Future Years Defense Program (FYDP), it is as a single column labeled "to complete." The PB also does not include much history, with most of it in a single column labeled "prior years." The SAR reports costs in both Then Year (TY) and Base Year (BY) dollars, while the PB reports exclusively TY dollars. The SARs are the Office of the Secretary of Defense (OSD)'s primary data source for analyzing MDAPs. This dataset is what analysts from many different organizations typically use, per the recommendation of the OUSD(AT&L) staff, who describe SAR data as "the official numbers."

---

3 Department of Defense, *Department of Defense Selected Acquisition Reports*, Release No. NR-106-16, March 24, 2016, accessed March 27, 2017, https://www.defense.gov/News/News-Releases/News-Release-View/Article/703529/department-of-defense-selected-acquisition-reports-sars.

4 Defense Acquisition University, Glossary of Defense Acquisition Acronyms and Terms, Defense Acquisition Management Information Retrieval (DAMIR) System, accessed March 28, 2017, https://dap.dau.mil/glossary/pages/1742.aspx.

5 The President's Budget and annual SAR submissions both generally come out in the second quarter of each fiscal year—typically February for the budget and a month later for the SAR. The years on a matching budget and SAR set differ by two, with the SAR having the earlier date because the budget is named for the year ahead and the December SAR is a snapshot of the program in the recent past. For example, in the second quarter of FY 2016, first the FY 2017 budget was released, followed by the December 2015 SAR.

1. Why SARs Matter

The SARs are not the dataset used most often for decision making inside the Department of Defense. When admirals, generals, or political appointees make large resource decisions, analysts assemble datasets to suit the needs of the decision maker. These analysts may pull data from non-public systems or use data calls to relevant offices. Because the decision makers have clout, they obtain the data they need to support their decisions. Why then do we care about the quality of data in the SARs?

The SARs matter for two reasons: triggering and research. What we call triggering is why the SARs were created. While OSD and the Congress both can assemble any dataset they want through their analytical arms, they do not know what questions to ask without something triggering them. The Services trigger investigations when they seek milestone authorities from OSD. OSD can also trigger analyses for program reviews based on the Service's annual submissions, such as the program objective memorandum. Only the SARs provide regular information at the program level. For example, no other annual submission can tell OSD or the Congress about the projected procurement costs for a program that is expected to leave the development stage in five years. Also, if a program is showing growth in Research, Development, Test, and Evaluation (RDT&E) costs but a decrease in procurement costs, only the SAR will show those offsetting. These are questions that should be answered on a regular basis and are by the SARs, but the quality of the answer depends on the quality of the data in the SARs.

Research on defense acquisition is continuously occurring. Such research is often conducted by people in government agencies, think tanks, and universities. In the past, acquisition research looking across programs has considered quantity of cost growth,[6] setting of production rates,[7] comparisons among different commodity types,[8] and many other subjects. This research helps the government, and SARs are the best source for doing comparisons across programs. While it is the nature of research that we cannot predict which research projects will yield fruitful results, we know that better quality data will yield better research results, which should therefore help the government be more efficient and capable.

---

6 David L. McNicol, "Cost Growth in Major Weapon Procurement Programs," IDA Paper P-3832 (Alexandria, VA: Institute for Defense Analyses, October 2004).

7 William P. Rogerson, "Excess capacity in weapons production: An empirical analysis," *Defence Economics* 2, No. 3 (1991): 235–49, doi: 10.1080/10430719108404695.

8 Jeffrey A. Drezner et al., *Are Ships Different? Policies and Procedures for the Acquisition of Ship Programs* (Santa Monica, CA: RAND Corporation, National Defense Research Institute, 2011), https://www.rand.org/pubs/monographs/MG991.html.

2. Nunn-McCurdy Breaches

Nunn-McCurdy (N-M) breaches are established by statute. If an MDAP sustains too much cost growth, a significant review takes place that generally leads to either changes in the program or, occasionally, termination. PMs generally want to avoid N-M breaches. Too much cost growth is defined in terms of Average Procurement Unit Cost (APUC) or Program Acquisition Unit Cost (PAUC).

APUC = Procurement Costs / Procurement Quantities

PAUC = Total Program Costs / (Procurement + RDT&E Quantities)

There are eight possible Nunn-McCurdy breaches, four for APUC and four for PAUC. The breach calculation is performed by measuring the percentage growth in APUC or PAUC. A *significant* N-M breach occurs if the average unit costs have increased by >15 percent of the current approved acquisition program baseline (APB) or >30 percent of the original APB. A *critical* breach occurs when the average unit costs have increased by at least 25 percent against the current APB or 50 percent against the original APB. The original APB is the APB that is established during the Milestone (MS) B decision (formerly Milestone II).[9]

Each SAR contains a unit cost report that compares the current APUC and PAUC estimates to the original APB and a second unit cost report comparing the estimates to the current APB if the current APB is not the same as the original one.

3. Subprograms

An MDAP's baseline may indicate that it has multiple subprograms to increase visibility into the program's activities. If so, each unit produced and each dollar spent is assigned to one of the subprograms. Subprograms have been used to distinguish variants of a system such as two similar but different missiles or to look at different parts of a system, such as engines and airframes. Each year, each subprogram has its APUC and PAUC calculated and compared to the baseline. According to the Nunn-McCurdy Act,[10] if any subprogram exceeds its thresholds, a Nunn-McCurdy breach is declared for the entire program, not just the subprogram that exceeded its baseline.

The popularity of subprograms has changed through the years, as can be seen in Figure 1. The total number of programs each year did not change much, but declaring subprograms

[9] John R. Hiller et al., "WSARA 2009: Joint Strike Fighter Root Cause Analysis," IDA Paper P-4612 (Alexandria, VA: Institute for Defense Analyses, June 2010).

[10] 10 U.S.C. § 2433, the Nunn-McCurdy Act. It was an amendment to the FY 1982 National Defense Authorization Act and can be found at https://www.congress.gov/bill/97th-congress/senate-bill/815.

became less common from 1998 to 2009, when a rebound started. It is not clear what has caused these changes.

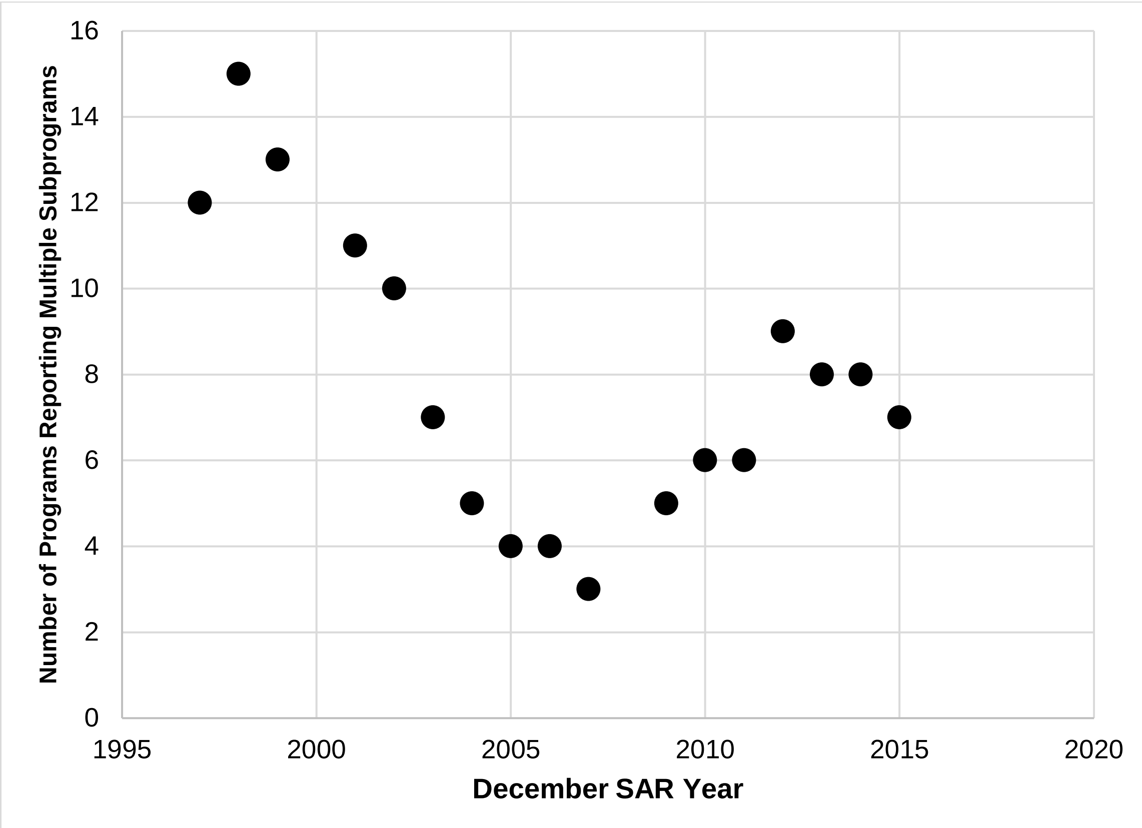

Figure 1. Subprograms in MDAPs 1997–2015

B. An Example of Budget and SAR Discrepancy: Gray Eagle

Quantity reporting in the SAR is the focus of this report. We begin with a few illustrative examples. The Army's MQ-1C Gray Eagle program acquires unmanned aerial systems (UASs). In the "Track to Budget" section of its 2015 SAR, the program identifies the following LINs within the "Aircraft Procurement, Army" appropriation:

- A00005 (MQ-1 UAS);
- A01001 (MQ-1 Payload, which includes funding for other programs); and
- A01005 (Common Sensor Payload Full Motion Video (CSP FMV), a sub-Line Item Number to A01001).

Both A01001 and A01005 are listed as "shared." The quantities and costs found in these LINs in PB 2017, however, differ from those in Gray Eagle's 2015 SAR, as shown in Table 1. Note that both sources project the program to finish in fiscal year (FY) 2018.

Table 1. Gray Eagle Program SAR and Budget Comparison

Data Source	FY 2015 Q	FY 2015 Cost (TY $)	FY 2016 Q	FY 2016 Cost (TY $)	FY 2017 Q	FY 2017 Cost (TY $)	FY 2018 Q	FY 2018 Cost (TY $)
PB 2017	19	$246,490 k	17	$355,445 k	0	$60,117 k	0	$10,806 k
Dec 2015 SAR	2	$246,400 k	3	$322,200 k	0	$60,200 k	0	$15,200 k
Difference	**17**	**$90 k**	**14**	**$33,245 k**	**0**	**-$83 k**	**0**	**-$4,394 k**

Note: k – thousand; Q – quantity.

The cost differences in FY 2015 and FY 2017 are minimal, but there is no obvious explanation for any of the difference, including the more significant differences in FY 2016 and FY 2018 costs. In PB 2017, the unit of accounting for this program is one unmanned airplane. While this is straightforward to understand, the capability is also dependent on how many ground assets for operating the systems are acquired and on the differences between aircraft, as they are not all the same. In the SAR, the quantity is measured in companies, each of which contains several aircraft with different configurations and some amount of ground equipment. While there is a standard measure for what a company is, not all companies fit the standard description. While the SAR does include a great deal of detail in various written sections, this makes it difficult to use the quantities in the data for quantitative analysis.

C. A Complex Example: The CH-47F Chinook Program

The Army's CH-47F Improved Cargo Helicopter program demonstrates challenges that can occur when counting quantities across years in both the PB and SAR. This program builds Chinook helicopters, which are easy to count, yet there are serious questions when looking at the data.

First, the CH-47F program's definition of one *unit* has changed over time. In the early days of the program (as reflected in the original June 1998 SAR), the plan was to SLEP[11] 300 existing CH-47D helicopters to an updated configuration, which would be called the CH-47F. In PB 2005, the plan was to SLEP 287 CH-47D helicopters to the CH-47F configuration, and 50 MH-47E Special Operations helicopters to a new MH-47G configuration. The definition of a unit had changed to include both CH-47D/F conversions and MH-47E/G conversions, which produce distinct end items and have different expected costs.

[11] SLEP is the acronym for "Service Life Extension Program," and is often used as a verb in defense circles. A SLEP can be funded with either procurement or O&M dollars.

The Army’s February 2007 budget justification forms expanded the set of planned activities to include all of the following activities:

- SLEPs of CH-47D to CH-47F
- SLEPs of MH-47E to MH-47G
- New builds of CH-47F from scratch for Active Duty Army units
- New builds of CH-47F in a different configuration for National Guard units

The reported and projected unit costs for these activities were all different. More to the point, the definition of a unit now included not only a remanufactured existing helicopter, but also a newly built helicopter of the same design. While these may be functionally identical from an operational point of view, there are reasons why an analyst would want to know how many of each were to be built—and at what cost. To further complicate matters, the helicopters produced (both SLEP and new build) employ a mix of mission subsystems, some of which could be repurposed from a remanufactured helicopter or other existing decommissioned helicopter, and some of which had to be built (and purchased) new. The type and number of repurposed subsystems continued to vary from year to year, so that the production inputs (and price) even for new-build Active Component CH-47Fs were different from year to year.

The end result of these changes is that any given unit produced by the CH-47F program might have any one of the MH-47G, CH-47F Army, or CH-47F National Guard configurations. A CH-47F unit might be remanufactured or built new. Whether remanufactured or new, it might include some unspecified mix of government-furnished (free) and contractor-furnished (at a price) mission subsystems. For example, as of the 2013 PB submission (February 2012), 43 new build units had been produced at an average cost of $15.0 million, of which $1.1 million per unit was for government-furnished equipment (GFE). The estimated cost to complete the new build program was $2.19 billion for 112 units, or $19.5 million per unit, of which $2.4 million per unit was expected to be GFE. This reflects the expectation that units authorized through FY 2013 would use recovered avionics suites from existing aircraft, but that half of the new build units after that would require new-build (contractor-furnished) avionics. There were clearly anticipated differences in components and cost between units produced up to that point and units expected to be produced in the future.

Furthermore, there are inconsistencies between the SAR and the PB submissions regarding which units comprise the CH-47F program. How new builds versus SLEPs are counted in different years is unusual and is described in detail in Section 4.A of this paper.

D. Organization of this Paper

We have divided the common differences among unit definitions into three buckets: changes over time, mixed types of units, and reporting accidents. It is not uncommon for more than one category to apply to a given program; the Chinook has all three. The next three chapters describe what each of these categories means, how confusions arise, and what analysts should do when trying to use cost reporting data. In Chapter 5, we make some modest recommendations for modifications to acquisition data reporting that could help make the data more useful for many sorts of analyses. As part of those recommendations, in Chapter 6 we consider the Joint Light Tactical Vehicle (JLTV) program, how its reporting might have been done differently, and what the ramifications of those differences might have been.

2. Changes over Time

It is implicit in the idea of a unit that every instance of the unit should be identical. Every inch should be the same length, every second should have the same duration, and every run scored in a baseball game should count equally. As noted above, this is often not true of procurement units in MDAPs. One reason that non-identical units might arise is that the product may evolve over time. Even when counting quantities is simple, such as when counting helicopters or ships, the units procured at different times are usually different in both cost and capability.

A. DDG-51

The Navy's DDG-51 program bought its first ship, *Arleigh Burke*, in 1985, and it is still buying ships today, with the December 2015 SAR listing the final buy as two ships in 2022. One might expect the cost of each ship to be about the same and a more sophisticated analyst might expect the cost of each successive ship to decrease due to learning. The data tell neither story.

In Figure 2, we see that the historic annual unit cost of these destroyers shows large jumps, steep falls, and long periods of gradual upward slopes. The usual interpretation of this pattern is that the jumps come from adopting new designs and the steep falls come from learning how to build them—a process that includes both learning by the shipyard's staff and also from investments in technology. The upward slopes that are visible in much of the data are probably from gradual upgrades to the ships. The projection forecasts that the Navy will stop enhancing the ships and let the learning effect continue to dominate for the balance of the program. History suggests that is unlikely, as the units are likely to change during that period.

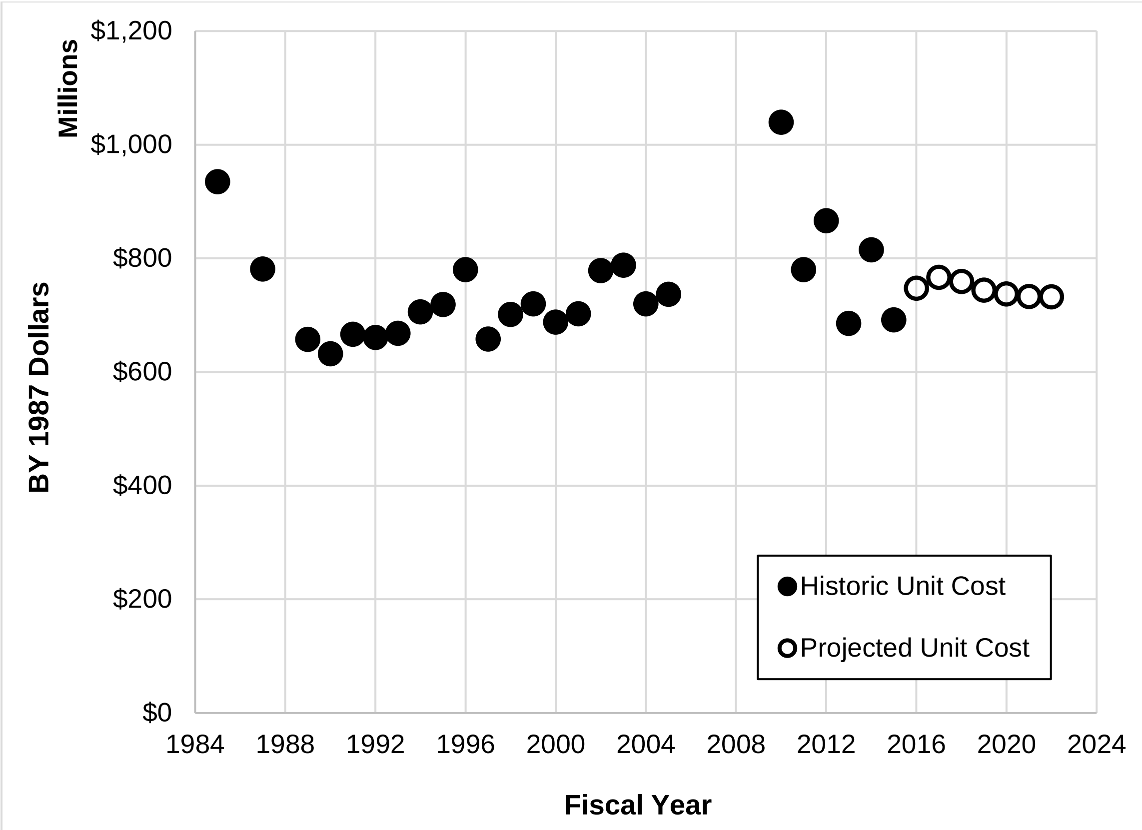

Figure 2. DDG 51 Unit Costs from the "Cost Quantity Information" Section of the December 2015 SAR

Figure 3 shows four ships of the *Arleigh Burke* class. The visible design differences among these ships accounts for some of the cost differences shown in Figure 2. *Fitzgerald*, on the left, has a bellmouth in the stern for a towed array sonar that is missing from *Sampson*, on the right. *Sampson* also has twin hangars for helicopters that *Fitzgerald* does not. In spite of these obvious differences, and numerous others, each of the four destroyers in the photo is considered one unit in all cost reporting. This issue is not unique to this program or even to shipbuilding; it is nearly ubiquitous in defense acquisition.

Source: "Destroyer Squadron 15, Japan Maritime Self-Defense Force Flex Their Combat Capability during Multi-Sail 2015," U.S. Pacific Command, March 25, 2015, accessed March 13, 2017, http://www.pacom.mil /Media/News/Article/581490/destroyer-squadron-15-japan-maritime-self-defense-force-flex-their-combat-capab/. (Photo by MC3 Raymond D. Diaz III)

Figure 3. 2015 PACOM News Photo of Four DDG-51 Class Destroyers in Guam

B. AIM-120 AMRAAM

The AIM-120 Advanced Medium Range Air-to-Air Missile (AMRAAM) program was established at a Defense Systems Acquisition Review Council MS I Review in November 1978. After an extended development period, an acquisition baseline of 24,320 units was set in December 1988. The first production units were authorized under the FY 1987 budget, and were fielded in 1991. The acquisition target was reduced to 16,427 missiles in a 1992 rebaselining that also doubled the expected per-unit cost.

The AIM-120 is still in production. The Air Force now intends to buy a total of 12,851 missiles, and the Navy an additional 4,461 missiles, for a total of 17,312. The final unit is projected to be authorized in FY 2025, almost 40 years after the first unit.

The explanation for the continued utility of the AIM-120 is that the missiles being produced today are nothing like the missiles that were produced in the early 1990s. Figure 4 shows the history of average unit cost by annual production lot for AMRAAM missiles. After a typical initial learning curve, it is clear that there have been major changes to the program over its history. In fact, there have been many upgrades, modifications, and

wholesale redesigns of the missile over time; the Teal Group reports seven.[12] Some were simply improvements, while others had new functions such as the AIM-120C3, designed with smaller control surfaces to fit inside the weapons bay of an F-22 Raptor, and the AIM-120D, which includes many new features such as GPS navigation and a two-way datalink.

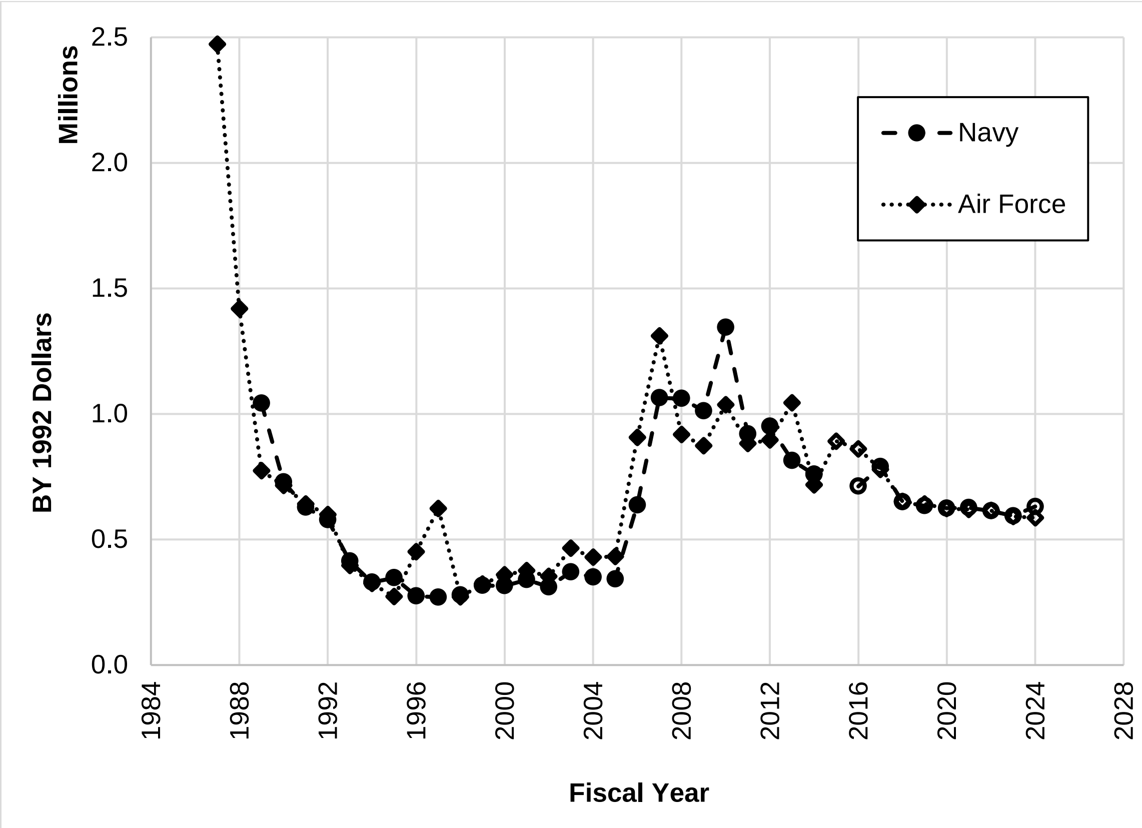

Figure 4. Annual Average Unit Cost for AMRAAM Missiles

There is no sense in which an AIM-120D is “the same thing” as an AIM-120A, or even an AIM-120C7. This is a clear instance in which the implicit assumption that units are interchangeable has been violated.

Of course, within the AMRAAM program there is no confusion about the kinds of missiles that are currently being produced, their capabilities, or plans for future improvements. The question, then, is how the program might adjust its data reporting to enhance transparency for planners, analysts, and oversight bodies.

---

[12] Teal Group Corporation, *World Missile Briefing*, December 2014, 133.

C. Tactical Aircraft

As part of an analysis of growth in aircraft prices, an IDA research team compiled data on the average empty weight of successive lots of various tactical aircraft programs.[13] The F-16 Falcon program filed its first SAR in 1975 and its last SAR in 1994. Over that period, it reported a single definition of *unit*, consisting of one aircraft. The average empty weight of the first production lot (funded in FY 1977) was 15,148 pounds. The average empty weight of the 1994 lot was 25 percent higher, at 18,971 pounds. Clearly, these were different aircraft. In fact, the Air Force refers to these aircraft using more than a dozen distinct designations (e.g., “F-16C Block 25”), and there are both single-seat and two-seat versions of many of the variants.

Similarly, the F-15 Eagle identified only one kind of unit as it grew 19 percent from 26,768 pounds in 1973 to 31,987 pounds in 1991, evolving from the single-seat F-15A Eagle to the two-seat F-15E Strike Eagle. Both the F-15 and the F-16 programs participated in the “Great Engine War”: some of each were built with Pratt and Whitney F-100 engines and others with General Electric F-110 engines. To a first approximation, no two production lots of either F-15 or F-16 used identical designs. Any analysis of program cost and/or schedule for these programs that does not account for the ongoing changes in the design of the aircraft will be misleading.

An alternative approach would be to declare a new program for modifications or upgrades that create a substantially different end item aircraft. This approach was taken by the Navy for the F/A-18. The F/A-18E/F aircraft was considered to be a sufficiently large departure from the predecessor F/A-18C/D that it was assigned a new program number and treated as a new start. Later, when the EA-18G electronic attack aircraft was developed as a variant of the F/A-18E/F, it too was designated a new start MDAP with separate accounting and units, despite sharing a production line with the F/A-18E/F.[14]

D. Tactical Vehicles

In 2014, IDA performed research on real price growth in the defense vehicle sector for the Director, Cost Assessment and Program Evaluation (CAPE).[15] In the course of that effort, it was found that the price of tactical military vehicles seemed to be growing at

---

[13] Stanley A. Horowitz, Bruce R. Harmon, and Daniel B. Levine, “Inflation Adjustments for Defense Acquisition,” *Defense and Peace Economics* 27, No. 2 (January 2016): 231–57, doi: 10.1080/10242694.2015.1093758.

[14] Similarly, the AIM-9 Sidewinder family of air-to-air missiles, unlike the AMRAAM program, has repeatedly spawned new-start MDAP programs. The AIM-9L (1977), AIM-9M (1982), AIM-9X (2000), and AIM-9X Block II (2011) programs were all treated as independent MDAPs with separate reporting and quantities.

[15] David M. Tate and Stanley A. Horowitz, “Hedonic Price Indices for Ground Vehicles,” IDA Document NS D-5467 (Alexandria, VA: Institute for Defense Analyses, May 2015).

roughly 15 percent per annum, far above either general inflation or other military systems. This was surprising; it seemed highly unlikely that prices for trucks would be rising much faster than (say) prices for fighter aircraft or naval vessels.

The paradox was resolved when the researchers discovered that there were significant year-over-year design changes in the vehicles in question. Even if the vehicle's designation (e.g., an M1025A2 HMMWV) did not change, there might be upgrades from one year's production to the next in armor, suspension, cooling system, or power train. These quality improvements, entirely undocumented in either SAR or PB submissions, accounted for roughly two thirds of the 15 percent annual price growth.

3. Mixed Types

Program offices often procure different end items at the same time. These items are usually similar to each other but substantially different. And yet for quantity reporting purposes, each is considered "one unit." This often comes about because of different missions or end users. Sometimes, the types are completely different.

A. Integrated Defensive Electronic Countermeasures

The Navy's Integrated Defensive Electronic Countermeasures (IDECM) program acquires electronics suites to protect the various F/A-18 aircraft from radio frequency guided missiles. IDECM achieved MS II approval in October 1995 as an Acquisition Category (ACAT) II program. Because of changes, it was designated as ACAT IC in March 2008 and issued its first SAR in June 2008. Below we include the Mission & Description section of the December 2015 SAR in its entirety.

> The Integrated Defensive Electronic Countermeasures (IDECM) System is a Radio Frequency (RF), self-protection electronic countermeasure suite on the F/A-18 aircraft. IDECM improves the survivability of the F/A-18 aircraft against RF guided threats during Air-to-Ground/Surface and Air-to-Air missions. The system is comprised of onboard components, which receive and process radar signals, along with onboard and offboard jammer components that transmit appropriate RF jamming responses.
>
> There are four IDECM variants in development, production, or sustainment. Blocks 1-3 are compatible with F/A-18E/F aircraft only. Block 4 is compatible with F/A-18C-F aircraft.
>
> IDECM Block 1: A federated suite, consisting of the ALQ-165 On-Board Jammer (OBJ) and ALE-50 expendable decoy.
>
> IDECM Block 2: An integrated suite, consisting of the ALQ-214 OBJ and ALE-50 expendable decoy.
>
> IDECM Block 3: An integrated suite, consisting of the ALQ-214 OBJ and ALE-55 Fiber Optic Towed Decoy.
>
> IDECM Block 4: A Hardware Engineering Change Proposal to the ALQ-214 OBJ to render it suitable for operation on F/A-18C/D aircraft, while retaining all functionality, when installed on F/A-18E/F.
>
> ALQ-214 Software Improvement Program (SWIP): ALQ-214 Software/Firmware updates that will enhance F/A-18 mission execution and improve mission survivability against modern air, land and naval threat systems by degrading (denying/delaying) threat ability to engage.

The SAR contains two subprograms: IDECM Blocks 2/3 and IDECM Block 4. The December 2015 SAR reports the Block 4 subprogram having an APUC of $2.502 million while the Block 2/3 subprogram has a far lower APUC of $0.090 million. This is because the quantities are so different. Block 4 has a quantity of 324, roughly the number of airplanes they will be protecting. Block 2/3 has a quantity of 12,805, although the Navy bought fewer than 600 F/A-18E/Fs in total. Eighty-five of the 12,805 were purchased with *1506* Navy Aircraft Procurement funds and the balance were or will be bought with *1508* Procurement of Ammunition, Navy and Marine Corps funds. We presume that those purchased with ammunition funding are only the disposable decoys. The unit costs based on the "End Item Recurring Flyaway" column in each year are presented in Figure 5.

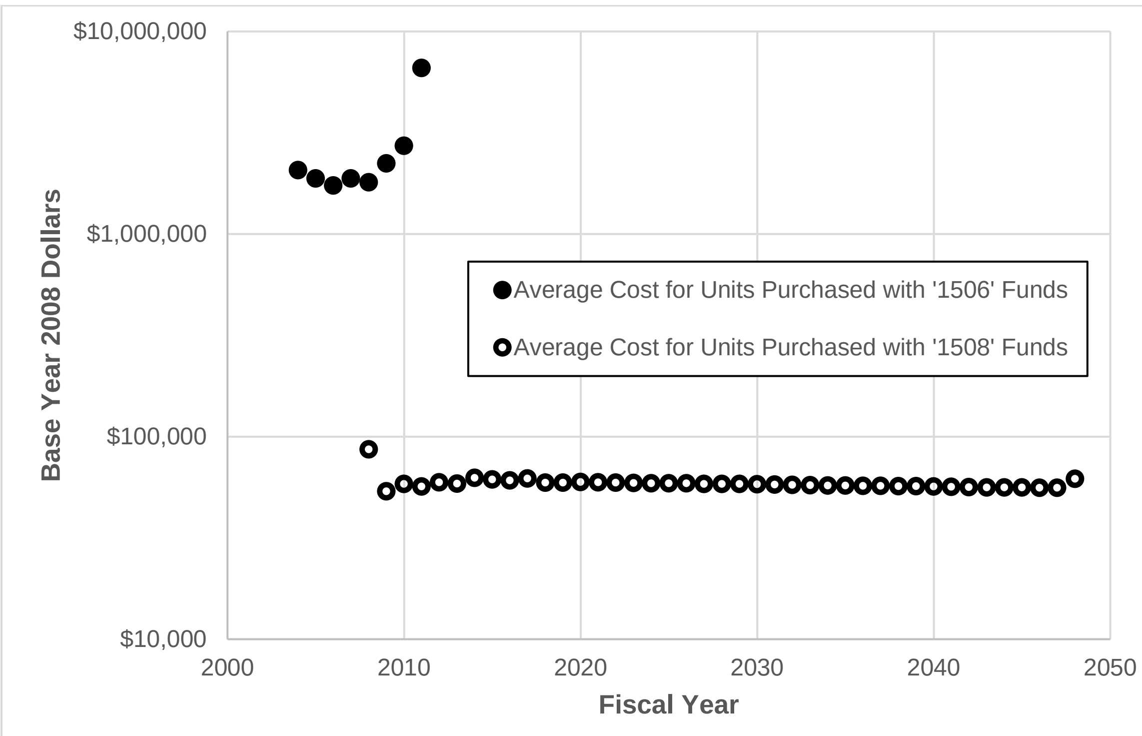

Figure 5. IDECM Block 2/3 Annual Unit Cost by Appropriation Type from the December 2015 SAR

Even though showing the two unit costs on the same chart requires plotting them on a logarithmic scale, the two are both considered *units* for the official unit cost calculation. Just within the more expensive *1506* units, it is clear that there have been significant changes, as the cost there does not follow a typical learning shape, which would be expected to slope down.

While the details have changed with time, the IDECM program has used this reporting system since it issued its first SAR in June 2008.

B. F-35

IDECM is difficult to quantify because it is a combination of an electronics suite with expendable decoys in different quantities. However, this pattern is also seen in cases where identifying units is much easier, such as airplanes.

From the December 2001 SAR through the December 2009 SAR, the F-35 (or Joint Strike Fighter (JSF)) program has reported one PAUC and one APUC for the program each year, even though the program has always been planning to acquire three different aircraft variants: F-35A conventional take-off and landing (CTOL), with the lowest price tag, generally the most capable once airborne; F-35B short take-off and vertical landing (STOVL) with reduced payload; and F-35C carrier variant (CV) for operating from aircraft carriers. The CTOL is for the Air Force, the STOVL is for the Marines, and the CV will be used by the Navy and possibly also the Marine Corps. Figure 6 shows the different variants the government plans to acquire.

Source: December 2015 F-35 SAR.

Figure 6. Joint Strike Fighter Variants

While the SAR's official unit cost reporting combines the three models in one average, the SAR does allow calculation of what the PAUC and APUC for the CTOL variant would be if it were tracked separately. This is possible because Navy and Air Force funding accounts are reported separately and the CTOL is the only version purchased with

Air Force funds. However, the CV and STOVL models are both purchased with *1506* Navy Aircraft Procurement funds and are not separated in the SAR at all, so the costs for those two variants cannot be segregated within the SAR data.

Starting with the FY 2011 PB request, the Navy's P-3a budget justification exhibits have separately reported costs and quantities for the two Department of the Navy variants, although all funding before that year places both variants together in the line for the CV variant. These exhibits do not show quantities or costs beyond the FYDP, and report only in TY dollars, making it impossible to compute the two variants' APUCs.

Beginning with the December 2010 SAR, the program was divided into two subprograms: aircraft and engine. The engine for the STOVL is significantly different than the other two because of the included lift fan, but the units purchased with Navy funds include some with the fan and some without it. Creating the subprograms did not clarify what the different F-35 variants cost.

4. Accidents

The confusions described above generally come about because of some decision by leadership about how the data should be presented;[16] this category, in contrast, is about cases in which it seems there were also outright errors in how the quantity numbers were put together. We do not know how frequent this is, but we know that it happens and can sometimes persist for several years. We do not suggest that any of the cases described below involve intent to confuse analysts, but they did have that effect.

This chapter presents three cases of accidents in SAR reports. The Chinook section is strictly a quantity accident, while the other two are more complicated and not strictly about quantity. We used the term “accidents” (as opposed to “mistakes” or “errors”) because it was the term a government official in AT&L applied to reporting anomalies for programs like Chinook. We show these three because they are the only three accidents of which we are aware. Accidents are inherently difficult to find; if they were obvious, they would be caught before publication.

A. CH-47F Chinook

As described in Section 1.C, the CH-47F Chinook Improved Cargo Helicopter program made a number of changes to its definition of *unit* over the course of the program. In the December 2015 SAR, however, the program apparently lost track of how it had been defining a unit and submitted quantity and cost forecasts that did not include all of the units identified in the simultaneous PB submission.

Figure 7 shows the discrepancy between predicted future quantities in the December 2015 SAR and the 2017 PB. Through FY 2017, the total quantities match perfectly. Beginning in 2018, units described as SLEP units in the PB are missing from the SAR forecast. As a result, the projected cost of these units is not included in the SAR calculations of APUC, PAUC, APUC growth, or PAUC growth.

---

16 Or, more precisely, leadership makes a decision about how the program should be managed and what systems it should produce, possibly without considering the impact this will have on the coherence and consistency of quantity or unit cost reporting.

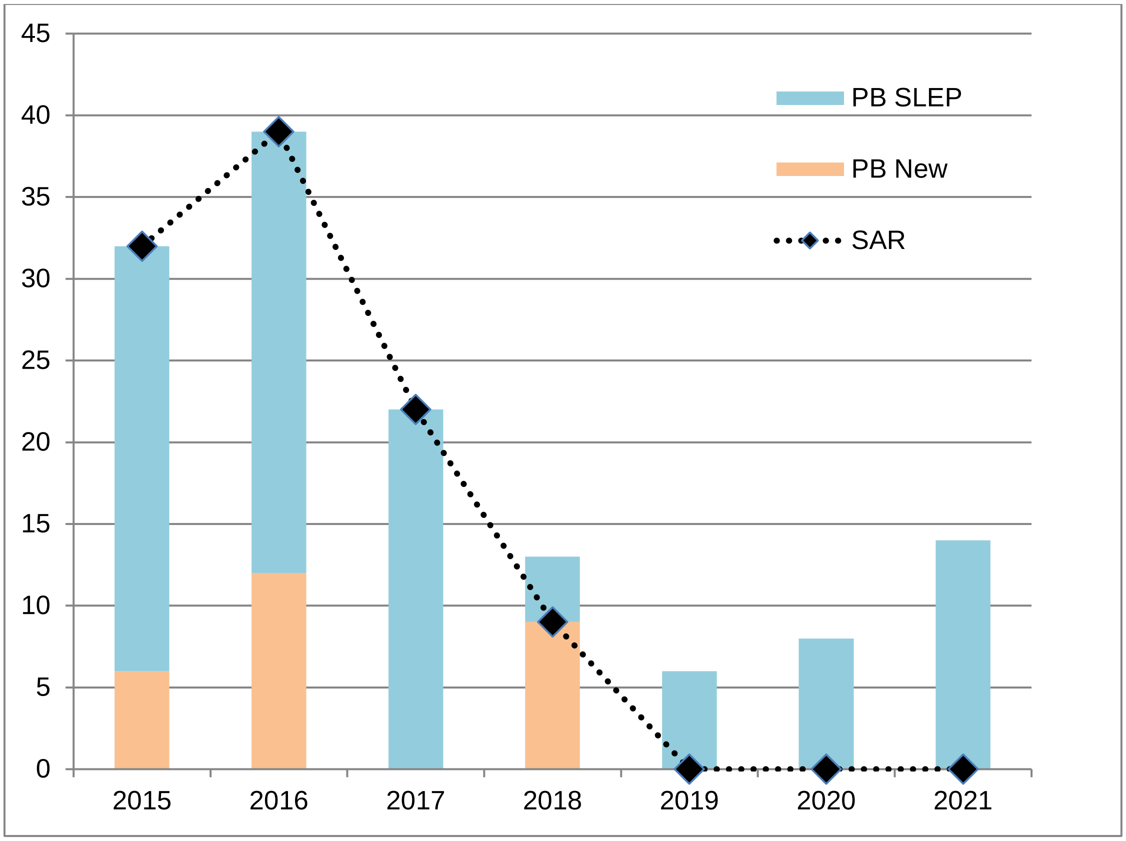

Figure 7. SAR versus PB Production Quantities for CH-47F

Since the SAR and the concurrent PB are required by law to agree on costs and quantities within the FYDP, this is clearly a mistake.

There is also an ongoing mismatch between the SAR and PB with regard to the past quantity produced. In the SAR every past unit produced is counted, regardless of whether it was a SLEP unit or a new build. In the PB, in the early years of the program there were no top-level quantities reported, presumably on the basis that upgrades to CH-47F configuration were just one of many ongoing upgrades in the Army's helicopter fleet. Typically, programs that perform multiple types of upgrade, but are not applying all of them to every legacy platform, report the number of each type of upgrade performed separately. They do not typically roll these up into a total quantity for the program's LIN, because the individual upgrades are not comparable and the number of platforms modified does not match the total number of any one type of modification.

When the decision was made to build CH-47F helicopters as new builds, the program began reporting a total quantity of units at the line item level, and chose to include both SLEP and new build units in this total. However, they never looked back to include previously produced SLEP units in the "prior quantity" total. As a result, each new SAR and PB submission disagree both on how many helicopters have been procured and on how many will have been procured in total when the program is finished.

B. ICBM Fuze Mod

Since 2006, SAR cost reporting standards require an annual breakout of program costs into four non-overlapping cost categories:

- End item recurring flyaway (EIRF) cost
- Non-end-item recurring flyaway (NEIRF) cost
- Nonrecurring flyaway cost
- Total support cost

EIRF cost includes direct costs associated with the production of specific end item units. NEIRF cost is used for fully configured production items that are not the quantity unit for this program, and are not produced one-for-one with the end item unit. For example, a remotely piloted aircraft program might report the cost of ground control stations as NEIRF. Similarly, a tactical vehicle program might report the cost of trailers or shelters as non-end-item. Nonrecurring costs are for capital costs that do not depend on the number of end items to be produced, such as production line tooling, testbed equipment, trainers/simulators, and shutdown costs. Support costs include systems engineering, program management, documentation, training, initial spares, and other program support activities. Cost analysts care about these categories because they can be expected to behave differently in response to program changes.

The Intercontinental Ballistic Missile (ICBM) Fuze Mod program is an Air Force program to replace the Mk21 fuze assemblies on W87 nuclear warheads. As such, this program would seem to have a very straightforward definition of quantity—namely, "number of fuzes replaced" (or alternatively, "number of warheads upgraded"). Since the fuze is a single assembly, and is the only product made by the program, there should be no non-end-item costs. It is surprising, therefore, that the December 2015 SAR for ICBM Fuze Mod categorizes more than a quarter of total program procurement costs as "non-end-item recurring flyaway." This must be a mistake. Given that the program reports zero anticipated nonrecurring costs, the most likely explanation is that anticipated nonrecurring costs were incorrectly reported as non-end-item.

As an aside, it is worth noting that the breakout of costs as recurring, nonrecurring, or support is required for both PB and SAR submissions—but the numbers in those respective reports rarely agree, even when the total costs reported match closely. This creates problems in understanding what is driving costs (and cost growth), and how costs might change under alternative acquisition plans.

C. Evolved Expendable Launch Vehicle (EELV)

The Evolved Expendable Launch Vehicle (EELV) program provides launch vehicles and launch services for US national security satellites. The program is run by the Air Force,

but it launches satellites for all three military services, as well as for the Intelligence Community (IC) and other federal agencies.

Like all SARs, the EELV SAR provides a “Track to Budget” section that identifies the funding accounts, PEs, and procurement line items that make up the program. In the case of EELV, 100 percent of SAR procurement costs are said to be funded under appropriation category *3021*, Space Procurement Air Force, line items MSEELC (Evolved Expendable Launch Capability) and MSEELV (Evolved Expendable Launch Vehicle). These accounts fund the ongoing provision of capable launch facilities and the individual launch vehicles used, respectively.

The December 2015 SAR forecasts 84 future launches, at a total cost of $37.6 billion, funded through appropriation *3021*, Space Procurement Air Force. This is incorrect in two different ways. First, many of those launches are not Air Force satellites, and will not be funded from that appropriation. This incorrectly characterizes the Air Force’s funding requirements in the FYDP and beyond, overstating them by tens of billions of dollars. It also potentially double-counts costs associated with Army, Navy, Department of Homeland Security, National Oceanic and Atmospheric Administration, and other non-IC satellites to the extent that those costs are correctly reported in those agencies’ own budget submissions.[17] Second, the SAR categorizes 100 percent of EELV procurement costs as EIRF, even though roughly half of the program’s costs arise though the “launch capability” line item, which is explicitly *not* aligned with specific launch vehicle units. Those costs should be categorized as nonrecurring or support, depending on the details of what they fund.

It is worth noting that the EELV program also has both changes over time and mixed type issues. The program uses multiple different launch vehicles, including variants of the Atlas V and Delta IV rockets, and neither rocket is the same today as it was when EELV began in 2000. EELV is also planning to use the SpaceX Falcon 9 rocket for some future launches.

[17] Some of the annual SARs give some breakdown on missions in the notes section, but the format is inconsistent and it is not part of the quantitative data.

5. Suggested Adjustments to Reporting

One possible response to the issues described above is to tell programs never to change what they are buying: once the baseline is set and the program is approved, the plan should be followed and the systems should not change. This assumption is implicit in the data reporting process. And yet, this has never been government practice and we do *not* recommend that it be adopted. Our military goes to great lengths to provide our warfighters with the best possible equipment and we should not forbid that just to make bookkeeping easier. We do offer some modest proposals that could make the reported data more useful, but first we need to be careful about incentives.

A. Data and Incentives

Data recording systems provide incentives. "You get what you pay for" is an old expression. In 2006, Dr. H. Thomas Johnson wrote: "Perhaps what you measure is what you get. More likely, what you measure is *all* you get."[18] If the acquisition system's data requirements are not aligned with the system's true goals, suboptimal performance may follow. This is exacerbated when penalties are associated with specific data reporting. Generally, people would prefer to report accurate data, but when the data will be used to justify punishment, they are incentivized to either change the facts that lead to the data—possibly in creative and unproductive ways—or to provide incorrect data.

The statute that defines the N-M breach specifies two measures (PAUC and APUC) and corresponding thresholds that influence program behavior. Since N-M reviews impose costs on programs (and are invariably unpleasant), and can trigger cancellation of a program, many people in defense acquisition, including PMs, try to avoid them. This likely accounts for some of what we see in data reporting today.

Any changes made to the system need to be considered in this light. If people's careers will depend on what data they report, then at times those data are more likely to reflect what is needed to satisfy the checker rather than reality. Furthermore, people will bend reality to make the data look "right" even if that will not yield the best actual result for national security.

---

[18] H. Thomas Johnson, "Lean Dilemma: Choose System Principles or Management Accounting Controls—Not Both," Chapter 1 in *Lean Accounting: Best Practices for Sustainable Integration*, ed. Joe Stenzel (New York: John Wiley & Sons, Inc., 2007).

B. Monitoring Changes over Time

If we accept that the units produced during the course of a program will change over time, PMs should be given useful and standardized ways to describe (and ideally quantify) those changes, both for past units produced and planned future production.

The current taxonomy of SAR Variance Categories recognizes the following possible reasons for cost growth:

- Economic – changes in then-year estimates due to revised escalation assumptions
- Quantity – changes in end item cost due to an increase or decrease in the number of units to be procured
- Schedule – changes in end item cost due to a revised production schedule
- Engineering – changes in end item cost due to alteration of the physical or functional characteristics of the end item
- Estimating – revision of earlier cost estimates based on new data or correction of identified errors
- Support – changes in any costs not categorized as flyaway costs,[19] regardless of cause
- Other – Exceptional circumstances, such as natural disaster or work stoppage. This category may only be used by permission of OUSD(AT&L).

Cost growth due to design changes should therefore always be categorized as Engineering, which lumps together planned and unplanned changes, as well as optional versus necessary changes. For oversight and analysis, it would be useful to be able to distinguish at least three sub-categories of "changes over time":

1. Pre-planned product improvements (P3I)
2. Unplanned changes (necessary and unnecessary)
3. Block upgrades or evolutionary acquisition

1. Pre-Planned Product Improvements (P3I)

P3I is a form of *spiral* acquisition, in which the first units produced do not include all of the capabilities that the procuring Service has identified as being required. The reasons for delaying might be budgetary, technical, operational, or some combination of those. The key is that the program has a plan from the beginning to add specific known improvements

[19] Or *rollaway* or *sailaway*, for vehicles and ships, respectively. Flyaway costs are defined in DoD 5000.4-M, *Cost Analysis Guidance and Procedures*.

in the future, and has developed cost and schedule estimates for those improvements. This allows P3I costs to be included in the SAR and other program submissions.

In the current SAR, or even the more detailed PB, it is difficult to report current or anticipated P3I costs in a transparent fashion. The additional costs beyond what the program would cost if the improvements were not made will be a mix of RDT&E costs (for developing and testing the new design), nonrecurring costs (for things like new documentation and tooling), EIRF costs (for actual production of the improved units), possible NEIRF costs (if improvements are made to non-end-item systems), and support costs (if the cost of support and/or spares for the new design is not exactly the same as for the original design).

In addition, the marginal costs of the improvements will generally be funded from the same RDT&E PE and/or procurement LIN as the original system. This means that there is no way to distinguish the P3I costs from the base system costs in the SAR, and no consistent way to report or forecast those costs in the PB.

For the marginal cost of improvements to be visible in the SAR, reporting would need to explicitly include P3I costs. One way to do this would be as follows:

1. If the planned improvements are small in number and to be done at a few discrete times during the production run, treat them like Block Upgrades (see Section 5.B.3).
2. If the planned upgrades are more numerous and continuous, establish a Planned Upgrades subprogram, and report the RDT&E and Procurement costs associated with planned changes to the original design under that subprogram. For each year in the SAR Annual Funding report, the program should report the following:
    a. Under the main end item subprogram, report the quantity produced or planned, and the estimated cost *if those units had been made to the original design*.
    b. Under the Planned Upgrades subprogram, report zero quantity, and the additional marginal procurement cost for the lot due to design changes. This additional cost should be split among EIRF, nonrecurring, and support costs in the usual way.
    c. Report RDT&E costs for the original design under the primary end item subprogram.
    d. Report RDT&E costs associated with planned design changes in the Planned Upgrades subprogram.

This system would allow analysts to clearly understand how much of the price change over time was driven by planned improvements and how much was unexpected. It would support meaningful learning curve modeling, and also provide some progress tracking of new capability insertions. The narrative portions of the SAR would describe the capability enhancements obtained to date, the plan for future insertion of new capabilities, and the unexpected changes made to the base program.

On the other hand, this system introduces a potentially onerous new type of reporting—namely, the hypothetical cost of the units if they had all been made to the original design. This is not information programs currently possess, and there are potential pitfalls and perverse incentives in how programs might choose to compute and report these counterfactual costs. In particular, cost growth due to design changes that might have been necessary in the base program (e.g., for safety reasons, to meet threshold requirements, or due to diminishing manufacturing sources) could be allocated either to the base subprogram or to the P3I subprogram, whichever seemed least likely to risk an N-M breach.

For N-M purposes, several regulatory changes might be beneficial. First, the primary end item and the Planned Upgrades should be treated as separate triggers. The primary end item would use the usual PAUC and APUC significant and critical thresholds. The Planned Upgrades subprogram might only have significant and critical limits based on total cost growth, or perhaps time-phased cost growth (e.g., average cost per year, rather than average cost per unit). Ideally, a breach on the Planned Upgrades subprogram would *not* imply a breach on the base subprogram (although the reverse would not be true).

Under this system, the main temptation for struggling programs would be to mischaracterize some of their core program cost growth as P3I, so as to avoid an N-M breach on the primary end item. By shedding planned improvements, the program could avoid having an N-M breach on either subprogram. This is not necessarily a bad thing. The oversight challenge would be to align operational test criteria with the phased capabilities to be produced.

2. Unplanned Changes

It is not uncommon for systems already in production to incorporate significant design changes that were not foreseen by the program. Reasons for this can include urgent operational needs from the field, correction of defects discovered post-fielding, implementation of Value Engineering proposals, or response to changes in the adversary/threat environment.

It is clearly unreasonable to require programs to report things they are not yet planning to do. For unplanned changes, the challenge is how to report them as they are discovered

and after the fact, in ways that transparently describe the reasons for any corresponding cost and schedule changes.

It would be ideal if SAR reporting of unplanned changes distinguished clearly between design changes driven by new performance requirements and changes required to meet the original program requirements. One possible way to accomplish this would be to add a new category, "Requirements," to the list of SAR variance categories. Cost changes due to design changes required to meet original program requirements (as of the current APB) would be classified as engineering variances. Cost changes due to new or modified performance requirements would be classified as requirements changes. For a program with a P3I subprogram, the base program and P3I subprogram would report separate cost variances, using the new category where appropriate.

Unfortunately, it is unlikely that programs would report these categories accurately. Not only are there strong incentives to categorize all cost growth as being due to new requirements, there is often genuine confusion within the Program Office about which requirements are part of the baseline and which have been added during the course of development and production. In theory, the Cost Analysis Requirements Document and other mandatory acquisition documents establish the baseline requirements assumed by the baseline cost estimate. In practice, this is not as clear, especially for programs that have been rebaselined at some point.

3. Block Upgrades, Evolutionary Acquisition, and Agile Development

Some programs know in advance that they intend to upgrade or replace the initial design with an improved future design, but do not yet know what those changes will be or what they will cost. They may not even know which attributes will be enhanced, since that decision will be based on feedback from actual mission use, changes in the threat environment, and/or the outcomes of uncertain technology development efforts. If multiple changes are made to the weapon system design at a few discrete points in time, these are often termed *block upgrades*. If many changes are made on an ongoing basis as their usefulness becomes known, this is sometimes referred to as *evolutionary acquisition*. In the special case of software programs doing repeated rapid insertion of new features in close collaboration with the users of the software, it is called *agile development*.

In each of these cases, the reporting challenge is that the program knows that they intend to spend money in the future, but they do not know what they will be spending it on, what it will cost, or when it will happen. The challenges for oversight and management are obvious—especially when a program being managed in this way is shoehorned into a reporting system designed for unchanging units. This is part of what happened to the RQ-4B Global Hawk program, which was intended from the beginning as an evolutionary acquisition, but was required to guess both content and schedule of future upgrades as part of its original acquisition baseline. Those guesses were then treated as firm requirements

by the acquisition system, even after Air Force leaders had changed their minds about both priorities and threshold performance.

In the case of block upgrades, one possibility is to simply declare a new program for each block. This is the approach taken by the AIM-9, AIM-9X, AIM-9X Block II missile programs; the F/A-18C/D, F/A-18E/F fighter aircraft programs; and the UH-60L, UH-60M Blackhawk helicopter programs (among many others).

Other programs have treated successive blocks as distinct official subprograms. This approach was taken by the Joint Air-to-Surface Standoff Missile (JASSM) program. The original program had no subprograms and developed the AGM-158 missile. During that development, the Air Force studied possible improvements to the missile, and decided to develop a second variant with longer range. The original AGM-158 was redesignated AGM-158A, and the new "JASSM-ER" (Extended Range) was designated AGM-158B. The program was split into two subprograms for reporting purposes, with JASSM-ER schedule, development costs, and production costs (and cost variances) reported separately. The Navy went even further with the new AGM-158C (LRASM) variant, deciding to make it a distinct program[20] rather than creating a new subprogram within the JASSM program. This may be because the new program is Navy-only, while JASSM is an Air Force program.[21]

An advantage of these approaches is that they isolate the unit cost of the new block from the past, rather than computing an average over all past blocks. It would defeat the purpose of the N-M legislation if 50 percent APUC growth in what is essentially a new weapon system was invisible because it was being averaged together with thousands of past units of completely different design.[22] A second advantage is that the block upgrade is clearly identifiable as design changes to meet new requirements, as opposed to design changes to overcome technical difficulties in achieving the original requirement.

One disadvantage of the subprogram approach, as currently implemented, is that an N-M breach by any block triggers a mandatory review of every subprogram. This seems counterproductive. Another reason for the JASSM program to split off a new program for the AGM-158C is that the Air Force would not want problems in a purely Navy subprogram to be able to trigger a breach that affects an Air Force program.

---

[20] PNO 449, "Offensive Anti-Surface Warfare Increment 1 (Long Range Anti-Ship Missile)," abbreviated as "OASuW Inc 1 (LRASM)".

[21] The "J" in JASSM stands for "joint" and at one point there was consideration of mounting this weapon on Navy aircraft. However, that has not happened, and all of the funds in the SAR are reported from Air Force appropriations.

[22] This is what has happened with the AIM-120 AMRAAM program, as described in Section 2.B above.

A disadvantage of both subprograms and separate programs is the difficulty of accounting for shared RDT&E, nonrecurring, and support costs, such as for testing equipment or software that is used by multiple blocks. For example, the RQ-4B Global Hawk family all use a common ground station. If this program had used separate subprograms for each distinct aircraft design, it would be inappropriate for the original RQ-4A subprogram to bear the cost of all upgrades to the ground station systems and software, given that all blocks benefit from those upgrades.

A logical response to this problem would be for the Global Hawk program to make the ground station systems a separate subprogram. The difficulty with this is that it would create the possibility of an N-M breach due to cost growth in a subprogram that accounts for only a small fraction of total program cost. A more reasonable approach would be for programs to be able to declare a single subprogram responsible for procurement of items other than end items. This subprogram would only be liable for an N-M breach if its estimated total cost (RDT&E + Procurement) grew to exceed a threshold percentage of the estimated PAUC for the overall program, which would require new legislation from the Congress.

C. Possible Methods for Handling Mixed Types

As the examples in Chapter 3 show, many solutions have been found to the mixed type problem, but all of them have drawbacks.

1. Subprograms

For some programs, subprograms have provided an elegant solution. For example, the Army's original Multiple Launch Rocket System (MLRS) program distinguished two subprograms: the mobile rocket launcher and the tactical rocket it would fire. This allowed the program to accurately track unit cost growth for both of the fully configured end items being developed and produced. The launcher was produced within its original cost estimate; the rocket experienced a critical N-M breach.[23] Similarly, the Army's PAC-3 suite of upgrades to the Patriot missile system was (after several schedule breaches in the first few years of development) divided into subprograms for the Missile Segment and the Fire Unit.

The fact that a unit cost breach in any subprogram triggers a breach in the whole program discourages their use, even where they seem like an obvious solution. A program without subprograms often has more leeway to do things that will make the cost growth look smaller. For example, if the MLRS program had not defined subprograms, but had treated the rockets as the end-item units, they would have shown a lower percentage cost

[23] Unfortunately, the program did not similarly distinguish the variant rockets being produced, or the later conversion of the entire system from an unguided rocket launcher to a guided missile launcher.

growth for the combined program than was seen for just the rocket subprogram. In addition, the program could have decided to produce fewer launchers than originally planned, reducing both PAUC and APUC without changing the official number of *units* being produced. Doing so might have avoided the N-M breach, at the cost of greatly reduced transparency regarding cost growth and reduced capability.

Making subprograms more appealing would require congressional action, possibly in an annual authorization bill, which seems possible if some way to maintain program cost accountability could be devised. The Congress might be willing to allow the Milestone Decision Authority to designate alternative triggers for programs with subprograms, especially if some of the subprograms involve far fewer dollars than others.

In theory, SAR reporting could be expanded so that each program could report simultaneously on multiple distinct end items without declaring subprograms. The principal distinction between this approach and subprograms would (presumably) be the mechanisms for deciding cost and schedule breaches. As with subprograms, it would be important in implementing this change to avoid creating perverse incentives to PMs. In particular, accurately defining multiple end items should not increase a program's chances of experiencing an N-M breach.

2. Multiple Programs

If a Service is planning to buy a mix of different end items in response to a given set of Mission Needs, they have some flexibility in deciding how to group those efforts into programs. It is not always obvious which grouping would best serve the needs of both the Service and the oversight community.

At one (unfortunate) extreme, the Army decided to make Future Combat Systems a single program with literally hundreds of different physical products ranging in size and complexity from light tanks down to man-portable UASs, along with many tens of millions of lines of software implementing communications, mission command, and networked fires. The official units for that program were Brigade Sets, of which 15 were to be produced. A prime "lead systems integrator" contract was awarded, with authority to reconsider the mix and capabilities of systems to be developed and procured in each Brigade Set. It is clear in hindsight that this offered no useful insight into the program's activities or progress.

At the other extreme, the Army decided to split procurement of their new AH-64E Apache helicopters into two separate programs, one for remanufactured aircraft and the other for new builds. A 2008 acquisition decision memorandum signed by the Army Acquisition Executive contains the following language.

> As a recently delegated Acquisition Category IC program, the AH-64E Apache program is comprised of two separate programs, the Remanufacture

> program and the New Build program. Each of these programs are separate and distinct with respect to the Acquisition Program Baselines (APB), and their funding lines; however, they have identical configurations and are produced on the same production line.[24]

This choice creates challenges for both the Army and OSD because it adds extra reviews and recordkeeping. Having multiple programs, as with subprograms, creates two triggers for an N-M breach, but it also means that any breach would affect only one of the two programs, whereas creating two subprograms would expose the entire program. It also splits what naturally feels like one program—indeed, the language above refers to it both as one program and as two in the same paragraph. Since both programs produce identical new AH-64E helicopters, why should they be separated? Although distinct for reporting purposes, they have common goals and management. They share a PM and a production contract,[25] but only the remanufacture program reports any RDT&E costs. The idea of an MDAP with no RDT&E costs is suspicious. Even within Apache, both programs list "Other Support" funds in their SARs, and since the two programs are producing identical helicopters, it is unclear how the Army decides whether a given support purchase will be credited to one program or the other. The bottom line is that you cannot understand what is going on in either program without considering the other, which would seem to violate the notion of what constitutes a program. Where there is only one distinct end item, having multiple programs is questionable.

Multiple programs should only be considered as an option in the case of block upgrades to an existing program (as discussed in Section 5.B.3), or when the set of things to be procured by a proposed new program involves the following:

- Significantly different product types with different acquisition risks,
- Multiple independent contracts with no real synergies, and
- Few significant interoperability requirements among systems.

In general, it will be rare that it is appropriate to split a new proposed program into multiple programs. An example of a program that perhaps should have been split is the Advanced Threat Infrared Countermeasures/Common Missile Warning System (ATIRCM/CMWS) program. This program originally intended to develop and field both a system to detect incoming anti-aircraft missiles (CMWS) and a system to respond to that alert by attempting to confuse the incoming missile (ATIRCM). Some aircraft were expected to carry both ATIRCM and CMWS systems while others just CMWS, although both types of installation

---

[24] Heidi Shyu, OASA(ALT), "Acquisition Decision Memorandum for the AH-64H Apache New Build Program," Memorandum for Program Executive Officer, Aviation, March 11, 2013.

[25] The December 2015 SAR for the remanufacture program lists four procurement contracts and two RDT&E contracts. The new build SAR only shows one contract, W58RGZ-12-C-0055, which is one of the four procurement contracts in the remanufacture program.

were initially called one unit. The two systems were somewhat independent in both applicability and effectiveness on different aircraft. No complex data interchange between the two systems was required, and the performance of neither system depended strongly on the performance of the other, except for the basic fact of successful detection by the CMWS systems. At a minimum, the two subsystems should have been reported as subprograms (as they were in their final SAR), but there was no technical reason to have combined the two into one program at all.

Another program that perhaps should have been split into multiple programs is the Stryker (originally "Interim Armored Vehicle") program. This program procured eight specialized variants of an existing non-developmental armored vehicle. Of these eight variants, six were relatively minor modifications of the existing design, while two[26] required extensive engineering changes to the original. An appropriate program management strategy would have been to make the six "minor modification" variants a single program (with six subprograms), and the two major redesigns either a second program with two subprograms, or two additional separate programs. That would have isolated the development risks of the two most risky projects from the more straightforward projects, and would have given OSD and the Congress better visibility of how the various projects were progressing. As it happened, the Stryker program experienced a significant (but not critical) N-M breach, driven entirely by problems in the two major redesign vehicles.

3. Different Cost Categories

Using the different cost categories in current SAR reporting can give some visibility into what is happening in a program, but generally does not allow better identification of different unit types. The distinction between "end items" and "non-end items" was not designed to capture differences among multiple distinct end items.

The Air Force's UAS MQ-9 Reaper program plans to procure 347 units, where each unit is an aircraft. The total procurement cost for the program is $9.2 billion BY 2008 dollars, but only 52 percent of that is EIRF. Another 22 percent is categorized as NEIRF, and the remaining 26 percent is Total Support. This information is useful for cost analysts, although this distribution has no impact on N-M reporting.[27] The aircraft quantity can be compared to the EIRF to understand those units, although there are no quantities reported for ground stations, so an analyst can only know what has been spent on them in total but

---

[26] The two were the M1128 Mobile Gun System (MGS), which attempted to mount a tank-like 105 mm direct fire cannon on a relatively light wheeled vehicle; and the M1135 Nuclear, Biological, Chemical, Reconnaissance Vehicle (NBC RV), which required a suite of sophisticated environmental sensors and a positive-overpressure internal environment.

[27] One could imagine the Air Force lowering the ratio of ground stations to aircraft not for operational reasons, but rather because they want to control APUC.

not what each costs. In this case, NEIRF is something like a subprogram for the ground stations, but it is less transparent than actual subprograms would be.

D. Reducing Accidents

When humans carry out activities, accidents are inevitable. Reducing accidents requires good processes. We have not analyzed the process for generating SARs or PB submissions. In principle, that could (and perhaps should) be done from a quality assurance point of view.

We were also told that AT&L/Acquisition Resource and Analysis (ARA) performs OSD’s checks on Service-submitted data and they do not have enough time to do it thoroughly. All of the draft SARs arrive at OSD in the same season. About a week after the data arrive, ARA meets with each program for about one hour, at which time ARA can ask questions. They feel that this process is insufficient and clearly there are changes that could reduce the accident rate.

The best way to improve ARA’s review is probably not only to add more time. While more time might help, ARA would probably also benefit from specialized tools to help them analyze the draft SAR data and quickly compare it to budget submissions, prior year SARs, and general rules about how acquisition programs typically behave. Proposing improvements to that process is beyond the scope of this report.

6. A Thought Experiment: JLTV

To illustrate the kind of reporting that would be necessary to improve both oversight and data utility for cost analysts, we looked at the Joint Light Tactical Vehicle (JLTV) program. JLTV makes a good example because it features planned changes, mix issues, and possible accidents as well. Seeing how JLTV reporting would need to change if it were to provide transparency regarding sources of cost growth—and in particular cost growth due to design changes over time—makes a good test of whether the approach is practicable or desirable.

A. Mix Issues

The JLTV program intends to procure four distinct variants of a light truck, intended to replace the High-Mobility Multipurpose Wheeled Vehicle (HMMWV, or "Humvee") in many operational roles for the Army and Marine Corps (USMC). The four variants, all common to both Services, are

- Utility vehicle,
- General-purpose vehicle (GP),
- Heavy guns carrier vehicle (HGC), and
- Close combat weapons carrier vehicle (CCWC).

PB submissions to date suggest that the most expensive variant, the CCWC, is expected to be 10 to 20 percent more expensive than the least expensive variant, the Utility vehicle. Nevertheless, APUC, PAUC, and affordability goals for the program are defined as total cost (Army plus USMC) divided by total quantity (Army plus USMC, all variants). The USMC buy is planned to finish within a few years, but the Army is currently planning to buy more than 49,000 units over 20+ years. Any changes in the total quantity to be purchased will be impossible to analyze without also knowing the revised mix of variants.

In addition, the JLTV program will be producing a variety of optional armor kits, customization packages (such as fording kits), trailers, and other support equipment not classified as part of a fully configured end item.[28]

[28] Andrew Feickert, "Joint Light Tactical Vehicle (JLTV): Background and Issues for Congress," CRS Report RS22942, Congressional Research Service, May 31, 2017, https://fas.org/sgp/crs/weapons/RS22942.pdf.

B. Planned Changes

Early SAR submissions from the JLTV program showed a typical (if somewhat optimistic) learning curve, showing steadily declining average unit cost in each lot over the entire production plan. The anticipated learning was just enough to bring the predicted APUC for the program below its affordability goal of $400,000 per vehicle.

To the delight of the Army and the program office, early production of JLTV proved to be significantly less expensive than originally expected. The effect on the SAR of this happy outcome was intriguing. Where before the program had predicted a steady lot-by-lot unit price decrease over decades of production, the current SAR shows an early decrease, followed by an uneven but persistent unit cost *increase* for more than a decade, before unit costs again begin to decline sharply (as shown in Figure 8).

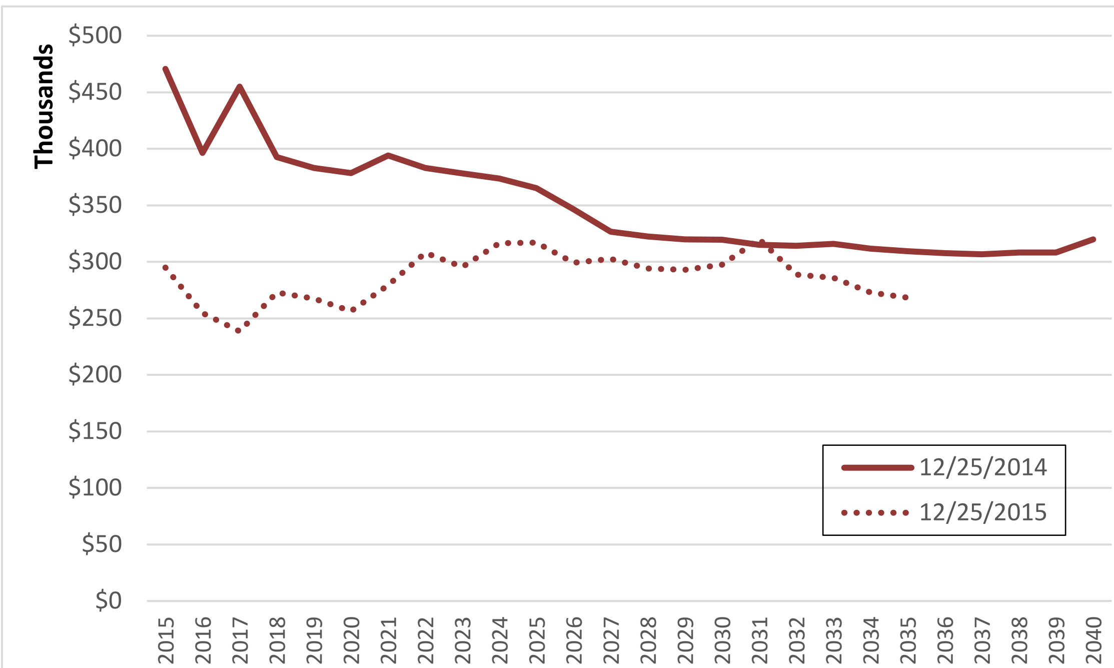

Figure 8. SAR End-Item Recurring Unit Cost Forecasts for JLTV

The most likely explanation for this pattern is that the program, given (in effect) extra funds to work with, has built a P3I effort into its cost estimates. Comparing the original learning projections against the new SAR forecast suggests that nearly half of the predicted unit cost of future JLTVs may go to pay for features that are not part of the current production designs.

C. Accidents

Because both the Army and Marine Corps are buying the same four variants of JLTV, and both have from time to time reported unit flyaway costs (actual and predicted) for each

variant in their PB submissions, it is possible to see that something strange is going on. The Army and Marines are not paying the same price for the same vehicle in the same year—not even close. The differences are not merely large; they are inconsistent, with the Army paying less in some years but more in others.

This unit price difference also shows up in the SAR, where Army and Navy budget appropriations are broken out, but SAR data alone do not tell how much of the difference in unit price might be explained by a difference in the mix of vehicle variants being procured by each Service, or whether there are large (and inconsistent) differences in the prices of the mission equipment packages used by each Service.

As a final anomaly, the last two PB submissions for JLTV have shown zero expected total nonrecurring costs over the life of the program. This cannot be correct; every materiel production program has nonrecurring tooling costs, documentation costs, data rights costs, etc. Indeed, the JLTV SAR forecasts more than $1.4 billion in nonrecurring costs, remaining fairly constant annually over the life of the program. This makes it impossible to use the justification books to reconcile the unit cost anomalies in the SAR.

D. A Transparent Reporting Structure for JLTV

For readers of the SAR to be able to understand and analyze unit cost growth over time in the JLTV program, a reporting structure similar to the following would be required. If such a structure were used, the Congress should strongly consider the legislative and regulatory adjustments to the N-M criteria proposed in Chapter 5.

1. Initial Subprograms

At program outset, the JLTV program was aware of four distinct variants, P3I initiatives, and non-end-item kits and support equipment to be produced. Even if the variants were expected to be fairly similar, with a high degree of commonality, this could change over the course of development (as was seen for the F-35 program, for example). Each variant should thus have its own subprogram.

In addition, there should be at least one subprogram for P3I if the P3I efforts are common to all variants; more than one if different planned improvements are intended for different variants. For illustration, we assume that there is one set of planned improvements that is common to all vehicles, and one set that only applies to the HGC and CCWC variants.

Finally, there should be a subprogram for support equipment. The complete initial set of subprograms is shown in Table 2.

Table 2. Suggested Initial JLTV Subprograms

Number	Subprogram	Description
1	Utility	Base design Utility Vehicle
2	General Purpose (GPV)	Base design GPV
3	Heavy Guns Carrier (HGC)	Base design HGC
4	Close Combat Weapons Carrier (CCWC)	Base design CCWC
5	P3I – Common	P3I common to all variants
6	P3I – HGC and CCWC	P3I specific to HGC and CCWC
7	Support equipment	Trailers, armor kits, etc.

2. Shared Costs

The program includes some costs that are common and came from the start; these are not P3I. These would include—but are not limited to—designing elements that are common to all versions, purchasing support equipment that could be used to service multiple variants, and non-recurring production costs. These must be split between the subprograms in some way—either evenly between the subprograms or on a weighted basis by the number to be procured. For example, common design elements should be split evenly between the variants that use them, while investments in production should be split based on the number of units to be purchased.

Another possibility would be to have a "Common Development" subprogram instead of the "P3I – Common" subprogram. This would make common development costs easy to track, but would lose the ability to see what part of the procurement of later vehicles was from the new improvements.

3. Additional Subprograms

Over the course of the program, the Army and/or Marines might decide to implement a major redesign of one or more variants as a block upgrade. In this case, a new subprogram should be created for each variant being redesigned. In the same way, if a new variant is developed in addition to the original four, that variant should get its own subprogram, and its own associated P3I subprogram if appropriate. Table 3 shows a possible future set of JLTV subprograms.

Table 3. Possible Future JLTV Subprograms

Number	Subprogram	Description
8	Utility Block II	Revised design Utility Vehicle
9	Light Scout (LS)	New design Light Scout Vehicle
10	P3I – Light Scout	P3I specific to new LS variant
11	Heavy Guns Carrier Block II	Revised design HGC

In this scheme, the base design subprograms and block upgrade subprograms would be subject to current N-M thresholds on APUC and PAUC growth. The P3I subprograms would be subject to a revised N-M threshold on total cost. The support subprogram would be subject to a new kind of N-M threshold on total support cost per end item produced under the base design subprograms (and any block upgrade subprograms). This would help to limit the chances of triggering an administratively painful and expensive delay for cost growth in a minor subprogram that does not materially affect the affordability of the program overall.

This is clearly a very complicated reporting structure relative to the current single program with no subprograms. It has the advantage of defining consistent units and showing what they cost. It also allows analysts and oversight bodies to distinguish cost growth in the original designs from cost growth in the modifications over time. When data were reported and stored on sheets of paper, this was likely too much. However, with today's databases, this kind of data would not be too burdensome and could be used in numerous ways.

7. Conclusions

The default assumption for any acquisition program is that all of the units it produces are identical and interchangeable. This is seldom true—consider asking an F-35A to land on a ship. Any analysis that assumes interchangeable units is making an unwarranted assumption that can lead to mistaken conclusions. The importance of these mistakes will vary, both with the details of the program and the nature of the analysis. We hope that this work can lead to two kinds of changes: one for analysts using acquisition data, and a second for policy makers defining reporting requirements for programs.

For analysts, the primary message is "Beware." It is not uncommon for invisible differences between units to be important to an analysis, as we saw with previous IDA studies of hedonic price indices for aircraft and tactical vehicles discussed in Chapter 2. Without additional data from non-SAR (and sometimes non-PB) sources, it is often impossible to understand the relationships among price, cost, and quantity in many programs. Such additional data is, unfortunately, not always available. Analysts need to know the limits of what can be inferred from the existing data.

For policy makers, there are many opportunities to improve data reporting requirements and guidance, and these come in three parts. First, there ought to be explicit acknowledgment that not all units are identical and some effort should be made to quantify unit-by-unit or lot-by-lot differences for analysis and oversight. Second, the rules need to encourage the desired behaviors. The current N-M rules are an excellent example of how rules incentivize behavior in ways that may be counterproductive. For example, IDECM's unit costs could be reduced by purchasing more towed decoys. When designing new reporting requirements, policy makers need to keep this in mind. Finally, the quality assurance processes applied to official data ought to be studied and improved. While some accidents are inevitable, the system today probably lets through more than it should. SARs are much like custom manufactured parts. Each one is unique, but good processes could still make them more uniform and useful.

Illustrations

Figures

Tables

References

10 U.S.C. § 2432: Selected Acquisition Reports. Accessed March 28, 2017. http://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title10-section2432&num=0&edition=prelim.

10 U.S.C. § 2433: Nunn-McCurdy Act. https://www.congress.gov/bill/97th-congress/senate-bill/815.

Benkard, C. Lanier. "Learning and Forgetting: The Dynamics of Aircraft Production." *American Economic Review* 90, No. 4 (September 2000): 1034–54. http://aida.econ.yale.edu/~lanierb/research/Learning_and_Forgetting_AER.pdf.

Defense Acquisition University. *Defense Acquisition Management Information Retrieval (DAMIR) System*. Glossary of Defense Acquisition Acronyms and Terms. Accessed March 28, 2017. https://dap.dau.mil/glossary/pages/1742.aspx.

Department of Defense 5000.4-M, *Cost Analysis Guidance and Procedures*.

Department of Defense. *Department of Defense Selected Acquisition Reports*. Release No. NR-106-16. March 24, 2016. Accessed March 27, 2017. https://www.defense.gov/News/News-Releases/News-Release-View/Article/703529/department-of-defense-selected-acquisition-reports-sars.

Drezner, Jeffrey A., Mark V. Arena, Megan McKernan, Robert Murphy, and Jessie Riposo. *Are Ships Different? Policies and Procedures for the Acquisition of Ship Programs*. Santa Monica, CA: RAND Corporation, National Defense Research Institute, 2011. https://www.rand.org/pubs/monographs/MG991.html.

Feickert, Andrew. "Joint Light Tactical Vehicle (JLTV): Background and Issues for Congress." CRS Report RS22942. Congressional Research Service, May 31, 2017. https://fas.org/sgp/crs/weapons/RS22942.pdf.

Harmon, Bruce R. "(U) Cost Estimating Techniques for Tactical Aircraft Manufacturing Labor." Unclassified/PI/LR. IDA Paper P-4490. Alexandria, VA: Institute for Defense Analyses, May 2010.

Hiller, John R., Scott A. Arnold, Ji S. Byun, Harley A. Cloud, Alexander O. Gallo, Matthew W. Gonwa, Bruce R. Harmon, Prashant R. Patel, Colin D. Sullivan, and Patricia F. Bronson. "WSARA 2009: Joint Strike Fighter Root Cause Analysis." IDA Paper P-4612. Alexandria, VA: Institute for Defense Analyses, June 2010.

Horowitz, Stanley A., Bruce R. Harmon, and Daniel B. Levine. "Inflation Adjustments for Defense Acquisition." *Defense and Peace Economics* 27, No. 2 (January 2016): 231–57. doi: 10.1080/10242694.2015.1093758.

Johnson, H. Thomas. "Lean Dilemma: Choose System Principles or Management Accounting Controls—Not Both." Chapter 1 in *Lean Accounting: Best Practices for*

Sustainable Integration, edited by Joe Stenzel. New York: John Wiley & Sons, Inc., 2007.

McNicol, David L. “Cost Growth in Major Weapon Procurement Programs.” IDA Paper P-3832. Alexandria, VA: Institute for Defense Analyses, October 2004.

McNicol, David L. “Post-Milestone B Funding Climate and Cost Growth in Major Defense Acquisition Programs.” IDA Paper P-8091. Alexandria, VA: Institute for Defense Analyses, March 2017.

Rogerson, William P. “Excess capacity in weapons production: An empirical analysis.” *Defence Economics* 2, No. 3 (1991): 235–49. doi: 10.1080/10430719108404695.

Shyu, Heidi. OASA(ALT). “Acquisition Decision Memorandum for the AH-64H Apache New Build Program.” Memorandum for Program Executive Officer, Aviation. March 11, 2013.

Tate, David M., and Stanley A. Horowitz. “Hedonic Price Indices for Ground Vehicles.” IDA Document NS D-5467. Alexandria, VA: Institute for Defense Analyses, May 2015.

Teal Group Corporation. *World Missile Briefing*. December 2014.

US Pacific Command. “Destroyer Squadron 15, Japan Maritime Self-Defense Force Flex Their Combat Capability during Multi-Sail 2015.” March 25, 2015. Accessed March 13, 2017. http://www.pacom.mil/Media/News/Article/581490 /destroyer-squadron-15-japan-maritime-self-defense-force-flex-their-combat-capab/.

Abbreviations

ACAT	Acquisition Category
AMRAAM	Advanced Medium Range Air-to-Air Missile
APB	Acquisition Program Baseline
APUC	Average Procurement Unit Cost
ARA	Acquisition Resource and Analysis
AT&L	Acquisition, Technology, and Logistics
ATIRCM	Advanced Threat Infrared Countermeasures
BY	Base Year
CAPE	Cost Assessment and Program Evaluation
CCWC	Close Combat Weapons Carrier
CMWS	Common Missile Warning System
CSP FMV	Common Sensor Payload Full Motion Video
CTOL	Conventional Take-Off and Landing
CV	Carrier Variant
D,PARCA	Director of Performance Assessments and Root Cause Analyses
DAMIR	Defense Acquisition Management Information Retrieval
DoD	Department of Defense
EELV	Evolved Expendable Launch Vehicle
EIRF	End Item Recurring Flyaway
FY	Fiscal Year
FYDP	Future Years Defense Program
GFE	Government-Furnished Equipment
GP	General Purpose
HGC	Heavy Guns Carrier
HMMWV	High-Mobility Multipurpose Wheeled Vehicle
IC	Intelligence Community
ICBM	Intercontinental Ballistic Missile
IDA	Institute for Defense Analyses
IDECM	Integrated Defensive Electronic Countermeasures
JASSM	Joint Air-to-Surface Standoff Missile
JASSM-ER	JASSM Extended Range

JLTV	Joint Light Tactical Vehicle
JSF	Joint Strike Fighter
k	Thousand
LIN	Line Item Number
LS	Light Scout
MDAP	Major Defense Acquisition Program
MGS	Mobile Gun System
MLRS	Multiple Launch Rocket System
MS	Milestone
MSEELC	Evolved Expendable Launch Capability
MSEELV	Evolved Expendable Launch Vehicle
NBC RV	Nuclear, Biological, Chemical Reconnaissance Vehicle
NEIRF	Non End Item Recurring Flyaway
N-M	Nunn-McCurdy
O&M	Operations and Maintenance
OASA(ALT)	Office of the Assistant Secretary of the Army for Acquisition, Logistics, and Technology
OBJ	On-Board Jammer
OSD	Office of the Secretary of Defense
OUSD(AT&L)	Office of the Under Secretary of Defense for Acquisition, Technology and Logistics
P3I	Pre-Planned Product Improvement
PARCA	Performance Assessment and Root Cause Analysis
PAUC	Program Acquisition Unit Cost
PB	President's Budget
PE	Program Element
PM	Program Manager
Q	Quantity
RDT&E	Research, Development, Test, and Evaluation
RF	Radio Frequency
SAR	Selected Acquisition Report
SLEP	Service Life Extension Program
STOVL	Short Take-Off and Vertical Landing
SWIP	Software Improvement Program
TY	Then Year
U.S.C.	United States Code

UAS	Unmanned Aircraft System
US	United States
USMC	United States Marine Corps

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
xx-12-2017	Final	May 2016 - May 2017

4. TITLE AND SUBTITLE
Complexity in an Unexpected Place: Quantities in Selected Acquisition Reports

5a. CONTRACT NUMBER HQ0034-14-D-0001
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Davis, Gregory, A.
Giles, Margaret, L.
Tate, David, M.

5d. PROJECT NUMBER
5e. TASK NUMBER AY-7-4166
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Defense Analyses
4850 Mark Center Drive
Alexandria, VA 22311-1882

8. PERFORMING ORGANIZATION REPORT NUMBER
IDA Paper P-8490
Log: H 17-000283

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Deputy Director for Program Assessments
Office of Performance Assessments and Root Cause Analyses
Office of the Assistant Secretary of Defense for Acquisition
5611 Columbia Pike
Falls Church, VA 22041

10. SPONSOR/MONITOR'S ACRONYM(S)
PARCA/OASD(A)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We have looked at the definition of units in numerous acquisition programs and discovered that the units reported are almost never simple; in some programs no two units are the same, and almost invariably the units produced at the end of a long production run are substantially different from the early ones. We have identified three reasons why the units may differ. The first reason is changes over time, generally as system capabilities are improved. The second is due to mixed types, where units that are inherently dissimilar, such as CH-47F and MH-47G helicopters, are produced by the same program and each is called one unit. The final reason why units can differ is reporting accidents. We give examples of all three and discuss possible methods of improving the reporting requirement.

15. SUBJECT TERMS
Acquisition; Data Reporting; Quantity Definition; Nunn-McCurdy

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Eggan, Peter
Unclassified	Unclassified	Unclassified	Same as Report	55	**19b. TELEPHONE NUMBER** *(Include area code)* (703) 697-3730

Reset